Rashid G. Alakbarov[1], Fahrad H. Pashayev[2], Mammad A. Hashimov[3] DOI: 10.25045/jpit.v07.i2.01
[1,3]Institute of Information Technology of ANAS, Baku, Azerbaijan
[2]Institute of Control Systems of ANAS, Baku, Azerbaijan
[1]rashid@iit.ab.az, [2]pasha.fahrad@gmail.com, [3]mhashimov@iit.ab.az

DEVELOPMENT OF THE METHOD OF DYNAMIC DISTRIBUTION OF USERS' DATA IN STORAGE DEVICES IN DATA PROCESSING CENTERS

The paper reviews a dynamic storage distribution based on cloud technology in data processing centers. The method is proposed for the dynamic distribution of the users' data among storage devices with various characteristics depending on the intensity of the data use. The method provides more efficient use of the storage resources by the users.

Keywords: *data processing center, cloud technology, cloud services, storage capacity, data call frequency, virtual resource, storage resource, simple and weighted moving average method.*

Introduction

At present, intensive investigations are being conducted for an effective use of the computing and storage resources of data processing centers with the help of Cloud Computing [1]. Cloud Computing is used for the development of distributed computing systems for the solution on high performance problems based on the computer network. Cloud Computing enables organizations to use computing and storage resources of the data processing centers more efficiently. With the help of this technology, the user data is stored and mined on the Cloud Computing servers, at the same time, the results are viewed through browsers. Cloud Computing service facilitates data processing centers to widely use the clustering and virtualization of computing and storage resources [2].

Cloud Computing enables to scale and use physical resources (e.g. processor, storage and disk space) through the Internet. In this case, the data processing and storage processes are considered as a type of service. The proposed model allows the data processing center to attract more users providing optimal distribution of available storage and system resources among the users [3].

Dynamic storage distribution among the users in data processing centers and the problem statement

It is known that SAN technology is widely used in the development of virtual memory resources in data processing centers. Whilst applying traditional SAN technology, the allocated resource is at full disposal of the user. In addition, when the user does not take full advantage of the resource, other users can not use this resource. As a result, the other half of the resource remains unused [4].

Hypervisor is used for the optimal resource allocation among the users, and the efficient use of the virtual memory [5]. This system helps to eliminate the shortcomings of traditional SAN technology.

The application of Hypervisor enables the new options for the control of virtual memory resources. These new functions include process automation, automatic recovery mechanism after the accidents, reserving, Thin Provisioning and Auto-Tiering, and so on. The use of both, Thin Provisioning and Auto-Tierng, technologies, at the same time, have a great impact on increasing the efficiency of virtual memory management and distribution [6].

Thin Provisioning in SAN systems allows the user to hold only the currently utilized volume of the resources [7]. The resources are used more efficiently in this way. As the allocated memory resource is fully used its quota is recovered from reserve resources. Thin Provisioning is widely used in the virtualization of the storage resources, and ensures the use of storage without wastes. This technique allows the resource allocated for any purpose to capture the storage as much as it is fully used. The resource is not backed up unless it is used, nor does fill any hard drive capacity. This technology provides an opportunity to use the centralized large storage in a planned way.

Therefore, the resources are allocated only as much as it is used without waste. Commercial side is beneficial for both, the user and the cloud provider. Thus, the user simply pays for the actual use of the resource rather than reserved one. Moreover, the provider is free from unnecessary expenses, such as the purchase of additional equipment, its installation, deployment; at the same time, it is able to offer the same service for a lower price, which leads to attracting more users to the services [3].

It is known that storage devices in data processing centers are based on memory devices with variety of properties (SSD, magnetic disk devices, etc.). The prices of such storage devices differ. Depending on which devices the users' data stored in, the payment for the resource rent is diverse. If the user uses very little amount of data stored in the data processing center over the year, the use of resources stored in magnetic disk device will be better. In this case, the expenses will be less. As the high-speed devices (SSD) costs much, the user, of course, spends more when using these devices.

The option of automated files (data) distribution by groups (types of memory devices) is one of the important components of the optimal allocation of resources and providing efficiency of their use. The option enables avoiding peak loading, and grouping the files by their characteristics. With the help of so called function Auto-Tiering the first layer of data center virtual storage includes high-speed SSD (Solid State Disks). The next layer uses magnetic disk devices Attachid serial SCSI (SAS), which runs less than SSD and cost less. The third layer uses lower speed serial ATA (SATA) devices, which costs less than above-mentioned devices (Figure 1) [8]. Taking into account the abovementioned, cloud-based deployment of the users' data is of great importance.

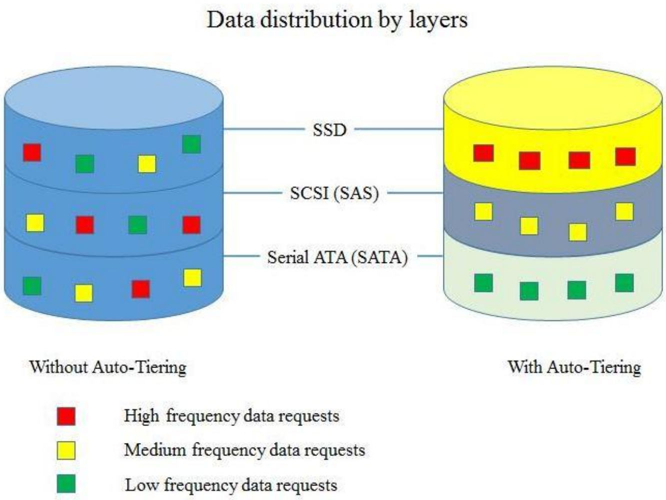

Fig.1. Deployment of storage devices by layers

Developing the algorithm and method for dynamic distribution of users' data among the layers of memory resources

The paper resolves the dynamic distribution of the files (data) among finite number of layers (memory devices).

Let's assume the followings to resolve the dynamic distribution of the files among finite number of layers:

- M - number of layers;
- V – total storage capacity of the layers;
- V_m- storage capacity of the m-th layer;
- I - number of files;

- F_i , - size of the file;
- N - number of observations;
- T_{in} , - moment of observation of the i-th file;
- f_{in} - current frequency of the calls to the i-th file at the n-th moment of observation;
- $\overline{f_i}$ - average frequency of the calls to the i-th file;
- h_{in} - current period of the calls to the i-th file at the n-th moment of observation;
- $\overline{h_i}$ - average period of the calls to the i-th file;

The storage capacity of the layers and the volumes of the files shall provide the following initial conditions:

1st condition: The sum of the storage capacity of the layers shall be equal to the total storage capacity V:

$$\sum_{m=1}^{M} V_m = V \, , m \in [1, M]; \tag{1}$$

2nd condition: The total volume of the files requested by the users shall not exceed 70 percent of the total storage capacity:

$$\sum_{i=1}^{I} F_i \leq 0.7 * V; \tag{2}$$

Note that current interval of the calls to the i-th file at the n-th moment of observation is calculated as follows:

$$h_{in} = T_{in} - T_{i(n-1)} \tag{3}$$

The relation between the current frequency of the calls to the i-th file at the n-th moment of observation f_{in} and the current interval h_{in} is as follows: $f_{in} = \frac{1}{h_{in}}$

The files' attributes and call methods are used to define call frequency to the files [9].

Preparation procedures include the regular arranging the calls frequencies to the files at any moment of observation. Therefore, they can be distributed among the layers. Let's assume that, $k_1, k_2, \dots, k_M$ are deployed in the layers as the result of distribution. The distribution shall provide the following conditions in accordance with initial conditions (1), (2):

$$\sum_{j=1}^{M} k_j = I; \tag{4}$$

$M_{m-1} = \sum_{j=1}^{m-1} k_j$ number of the files can be deployed in the m-1 number of layers up to m-th layer. k_mm number of files deployed in the m-th layer provide condition

$$\sum_{l=1}^{m_i} F_{M_{i-1}+l} \leq 0{,}7 V_m \tag{5}$$

in accordance with the second initial layer (2).

Consecutively, this distribution resolves the clustering of the files by the layers. The software implementing the distribution runs on the deployment in any m-th layer starting from the file, the serial number of which is $M_{m-1} + 1$, unless the condition (5) breaks up.

Though the method is simple, the date of calls to the files is not taken into account. The frequencies formed during the current calls may be random. Therefore, we must take into account frequencies formed during the previous observations, as well. The problem can be solved with the application of various version of Moving Average methods [10–12].

Simple Moving Average method: Since the calls frequency to the i-th file at the n-th moment of observation f_{in} , if we calculate the average value as the k number of elements, then

$$SMA_k = \frac{f_{in} + f_{i(n-1)} + \dots + f_{i(n-(k-1))}}{k} \tag{6}$$

Here SMA_k denotes Simple Moving Average. SMA_k shall be taken for the clustering instead of f_{in}.

Weighted Moving Average method: Applying Simple Moving Average each of the values of k number of frequencies calculated at various moments of observation will be of the same strength. Sometimes, the preference of the latter values may be significant. Therefore, the following formula of weighted moving average can be applied:

$$VMA_k = \frac{k * f_{in} + (k-1) * f_{i(n-1)} + \cdots + f_{i(n-(k-1))}}{k + (k-1) + \cdots + 1} \tag{7}$$

Here VMA_k denotes Weighted Moving Average and the latest values of the frequencies have higher weight. Correspondingly, VMA_k shall be taken for the clustering instead of f_{in} .

Exponential smoothing method: in exponential smoothing, taking any smoothing parameter α (0 <α <1) the sought parameter can be calculated as follows:

$$AR_k = a * \ f_{in} + a(1-a)f_{i(n-1)} + a(1-a)^2 f_i(n-2) + \cdots \tag{8}$$

Formula (8) shows that getting closer to 1 the latter values of parameter α are preferred more. In practical applications, the values of the parameter α varies between 0.01 and 0.30.

Thus, we may include one of the values, which are obtained through three various methods of call frequencies to the files, to the beginning of the algorithm of regular sequence:

- Current values. In this case, random changes of the current values can cause any file shift between the layers and imbalance of the process if the new value is not satisfactory.
- The values defined through Simple Moving Average method. The random changes in the current values do not cause significant balance disorders. The file is slowly taking its place in the layer in accordance with its frequency if the new value of the frequency is satisfactory.
- The values defined through Weighted Moving Average method. The random changes in the current values do not cause significant balance disorders. The file is taking its place more quickly than the Simple Moving Average method in the layer in accordance with its frequency if the new value of the frequency is satisfactory.
- The values defined through exponential smoothing method. In this case, current values are smoothed and the random changes in the values do not influence the final results. If the new values of the frequency are approved they will be placed in a certain layers quite quickly.

Finally, the followings shall be implemented for the solution of the dynamic distribution of the files among finite number of layers:

1. The call periods and frequencies to the experimental data files at the moments of observation shall be formed;
2. The average frequencies shall be formed out of Current values, Simple Moving Average or Weighted Moving Average for clustering;
3. The average frequencies and appropriate files shall be arranged in increasing order regularly;
4. Clustering and distribution by groups shall be carried out if condition (5) is provided.

Conclusion

The article analyzes data call frequency of the users and consequently, proposes to use memory devices appropriate to the data call frequency. At the same time, dynamic distribution of files (data) between the finite number of layers (memory devices) M. The clustering method among the layers is provided using current values formed at the moments of observations of the calls to the files, and the values found through Simple Moving Average or Weighted Moving Average, and Exponential smoothing. Simple Moving Average or Weighted Moving Average techniques proposed in the article can be used to solve analogical problems.

Thankful

This study has been carried out with the financial support of the Science Development Fund under the President of the Republic of Azerbaijan **- Grant № EIF-2014-9 (24) KETPL-14/02/1.**

References

1. Alguliev R.M., Alakbarov R.G. Cloud Computing: Modern State, Problems and Prospects / Telecommunications and Radio Engineering, USA, 2013, vol.72, no.3, pp.255–266.
2. Marios D. Dikaiakos, George Pallis, Dimitrios Katsaros, Pankaj Mehra, Athena Vakali. Cloud Computing -Distributed Internet Computing for IT and Scientific Research / IEEE Internet Computing, 2009, № 9, pp.10–13.
3. Alakbarov R.G., Pashaev F.H., Hashimov M.A. Development of the Model of Dynamic Storage Distribution in Data Processing Centers / Information Technology and Computer Science, 2015, no.5, pp.18–24.
4. An introduction to thin provisioning. http://searchstorage.techtarget.com
5. Storage hypervisor tutorial. http://searchvirtualstorage.techtarget.com
6. Jeremy LeBlanc, Adam Mendoza, Mike McNamara, NetApp Ashish Yajnik, Rishi Manocha, Symantec. Maximize Storage Efficiency with NetApp Thin Provisioning and Symantec Thin Reclamation. September 2010, WP-7111.
7. Floyer David. Thin provisioning. http://wikibon.org/wiki/Thin provisioning
8. Automated Storage Tiering. www.datacore.com
9. Tanenbaum A.S. Modern Operating Systems. Second edition. Piter, 2007, 1038 p.
10. Anatoly Homonenko and others. Delphi 7. St. Petersburg,"BHV Petersburg", 2008, p.1200.
11. Thomopoulos N.T. Applied Forecasting Methods. Prentice-Hall, New Jersey, 1980, 369 p.
12. Hyndman R.J., Athanasopouls G. Forecasting: principles and practice. www.otexst.org/book/fpp